

PCT

- WORLD INTELLECTUAL PROPERTY ORGANIZATION  
International Bureau

## INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification <sup>7</sup> : <b>C12Q 1/68</b>	<b>A2</b>	(11) International Publication Number: <b>WO 00/31303</b> (43) International Publication Date: 2 June 2000 (02.06.00)
(21) International Application Number: PCT/US99/27732 (22) International Filing Date: 22 November 1999 (22.11.99) (30) Priority Data: 09/198,083                      23 November 1998 (23.11.98)    US (71) Applicant: EXACT LABORATORIES, INC. [US/US]; 63 Great Road, Maynard, MA 01754 (US). (72) Inventors: SHUBER, Anthony, P.; 11 Grant Street, Milford, MA 01757 (US). LAPIDUS, Stanley, N.; 12 Old Evergreen Road, Bedford, NH 03110 (US). RADCLIFFE, Gail, E.; 11 Palisades Street, Worcester, MA 01604 (US). (74) Agent: WILSON, Daniel, A.; Testa, Hurwitz & Thibault, LLP, High Street Tower, 125 High Street, Boston, MA 02110 (US).		(81) Designated States: AU, CA, JP, European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).  <b>Published</b> <i>Without international search report and to be republished          upon receipt of that report.</i>
(54) Title: METHODS FOR STOOL SAMPLE PREPARATION  (57) Abstract  <p>The present invention provides methods for the preparation of stool samples to increase the yield of relevant DNA, and further provides methods for isolating and analyzing target DNA for characteristics indicative of colorectal cancer. Methods for screening patients for the presence of cancerous or pre-cancerous colorectal lesions are also provided.</p>		

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

FOR THE PURPOSES OF INFORMATION ONLY

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TC	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	CN	Guinea	MK	The former Yugoslav	TM	Turkmenistan
BF	Burkina Faso	GR	Greece	ML	Mali	TR	Turkey
BG	Bulgaria	HU	Hungary	MR	Mauritania	UA	Ukraine
BJ	Benin	IE	Ireland	MN	Mongolia	UG	Uganda
BR	Brazil	IL	Israel	MR	Mauritania	US	United States of America
BY	Belarus	IS	Iceland	MW	Malawi	UZ	Uzbekistan
CA	Canada	IT	Italy	MX	Mexico	VN	Viet Nam
CF	Central African Republic	JP	Japan	NE	Niger	YU	Yugoslavia
CG	Congo	KE	Kenya	NL	Netherlands	ZW	Zimbabwe
CH	Switzerland	KG	Kyrgyzstan	NO	Norway		
CI	Cote d'Ivoire	KP	Democratic People's	NZ	New Zealand		
CM	Cameroon	KR	Republic of Korea	PL	Poland		
CN	China	KZ	Kazakhstan	PT	Portugal		
CU	Cuba	LC	Saint Lucia	RO	Romania		
CZ	Czech Republic	LI	Liechtenstein	RU	Russian Federation		
DE	Germany	LK	Sri Lanka	SE	Sweden		
DK	Denmark	LR	Liberia	SG	Singapore		

## METHODS FOR STOOL SAMPLE PREPARATION

### RELATED APPLICATIONS

This application is a continuation-in-part of U.S. Application Serial No. 08/876,638, filed June 15, 1997.

### FIELD OF THE INVENTION

This invention relates to methods for the early detection of colon cancer in  
5 patients, and more particularly to methods for preparing stool samples in order to increase the yield of nucleic acids.

### BACKGROUND OF THE INVENTION

Stool samples frequently must be prepared for medical diagnostic analysis. Stool samples may be analyzed for diagnosis of medical conditions ranging from  
10 parasitic, bacterial or viral infections to inflammatory bowel disease and colorectal cancer.

Colorectal cancer is a leading cause of death in Western society. However, if diagnosed early, it may be treated effectively by removal of the cancerous tissue. Colorectal cancers originate in the colorectal epithelium and typically are not  
15 extensively vascularized (and therefore not invasive) during the early stages of development. Colorectal cancer is thought to result from the clonal expansion of a single mutant cell in the epithelial lining of the colon or rectum. The transition to a highly vascularized, invasive and ultimately metastatic cancer which spreads throughout the body commonly takes ten years or longer. If the cancer is detected prior  
20 to invasion, surgical removal of the cancerous tissue is an effective cure. However, colorectal cancer is often detected only upon manifestation of clinical symptoms, such as pain and black tarry stool. Generally, such symptoms are present only when the disease is well established, and often after metastasis has occurred. Early detection of colorectal cancer therefore is important in order to significantly reduce its morbidity.

- 2 -

Invasive diagnostic methods such as endoscopic examination allow for direct visual identification, removal, and biopsy of potentially cancerous growths. Endoscopy is expensive, uncomfortable, inherently risky, and therefore not a practical tool for screening populations to identify those with colorectal cancer. Non-invasive analysis of stool samples for characteristics indicative of the presence of colorectal cancer or precancer is a preferred alternative for early diagnosis, but no known diagnostic method is available which reliably achieves this goal.

Current non-invasive screening methods involve assaying stool samples for the presence of fecal occult blood or for elevated levels of carcinoembryonic antigen, both of which are suggestive of the presence of colorectal cancer. Additionally, recent developments in molecular biology provide methods of great potential for detecting the presence of a range of DNA mutations or alterations indicative of colorectal cancer. The presence of such mutations can be detected in DNA found in stool samples during various stages of colorectal cancer. However, stool comprises cells and cellular debris from the patient, from microorganisms, and from food, resulting in a heterogeneous population of cells. This makes detection of small, specific subpopulations difficult to detect reliably.

Use of the polymerase chain reaction (PCR) has made detection of nucleic acids more routine, but any PCR is limited by the amount of DNA present in a sample. A minimum amount of material must be present for specific analysis and this limitation becomes more relevant when one seeks to detect a nucleic acid that is present in a sample in small proportion relative to other nucleic acids in the sample, which is often the case when analyzing stool sample for detecting DNA characteristics of colorectal cancer. If a low-frequency mutant strand is not amplified in the first few rounds of PCR, any signal obtained from the mutant strand in later rounds will be obscured by background or by competing signal from amplification of ubiquitous wild-type strand.

An additional problem encountered in preparation of stool sample for detection of colorectal cancer is the difficulty of extracting sufficient quantities of relevant DNA from the stool. Stool samples routinely contain cell debris, enzymes, bacteria (and associated nucleic acids), and various other compounds that can interfere with traditional DNA extraction procedures and reduce DNA yield. Furthermore, DNA in

- 3 -

stool often appears digested or partially digested, which can reduce the efficiency of extraction methods.

## SUMMARY OF THE INVENTION

It has now been appreciated that the yield of nucleic acid from a stool sample is increased by providing an optimal ratio of solvent volume to stool mass in the sample. Accordingly, the invention provides stool sample preparation protocols for increasing sample nucleic acid yield.

In a preferred embodiment, methods of the invention comprise homogenizing a representative stool sample in a solvent in order to form a homogenized sample mixture having a solvent volume to stool mass ratio of at least 5:1, then enriching the homogenized sample for the target (human) DNA. The human DNA may then be analyzed for the characteristics of disease. Providing an optimal solvent volume to stool mass ratio increases the yield of nucleic acid obtained from the sample. An especially-preferred ratio of solvent volume to stool mass is between about 10:1 and about 30:1, more preferably from about 10:1 to about 20:1, and most preferably 10:1.

A preferred solvent for preparing stool samples according to the invention is a physiologically-compatible buffer such as a buffer comprising Tris-EDTA-NaCl. A preferred buffer is a Tris-EDTA-NaCl buffer comprising about 50 to about 100 mM Tris, about 10 to about 20 mM EDTA, and about 5 to about 15 mM NaCl at about pH 9.0. A particularly preferred buffer is 50 mM Tris, 16 mM EDTA and 10 mM NaCl at pH 9.0. Another preferred solvent is guanidine isothiocyanate (GITC). A preferred GITC buffer has a concentration of about 1 M to about 5 M. A particularly preferred GITC buffer has a concentration of about 3 M.

Also in a preferred embodiment, methods further comprise the step of enriching the homogenized sample mixture for human DNA by, for example, using sequence-specific nucleic acid probes hybridizing to target human DNA.

In an alternative preferred embodiment, the methods of the invention comprise homogenizing a stool sample in a physiologically-acceptable solvent for DNA in order to form a homogenized sample mixture having a solvent volume to stool mass ratio of at least 5:1; ensuring that the homogenized sample has at least a minimum number N of total DNA molecules to facilitate detection of a low-frequency target DNA molecule; and

- 4 -

analyzing the target DNA for the characteristics of disease, preferably by amplifying the target DNA with a polymerase chain reaction.

In another embodiment, the present invention provides methods for analyzing DNA extracted from stool which comprise homogenizing a stool sample in a solvent for  
5 DNA in order to form a homogenized sample mixture having a solvent volume to stool mass ratio of at least 5:1; enriching the homogenized sample for human DNA; ensuring that the enriched homogenized sample has at least a minimum number N of total DNA molecules to provide for detection of a low-frequency target DNA molecule; and analyzing the target DNA for DNA characteristics indicative of disease.

10 Methods of the invention are useful to screen for the presence in a stool sample of nucleic acids indicative of colorectal cancer. Such methods comprise obtaining a representative stool sample (i.e., at least a cross-section); homogenizing the sample in a solvent having a solvent volume to stool mass ratio of at least 5:1; enriching the sample for target human DNA; and analyzing the DNA for characteristics of colorectal  
15 cancer. Various methods of analysis of DNA characteristics exist, such as those disclosed in co-owned, copending U.S. Patent application Serial No. 08/700,583, incorporated by reference herein.

Methods of the invention also comprise obtaining a representative (i.e., cross-sectional) sample of stool and homogenizing the stool in a buffer, such as a buffer  
20 comprising a detergent and a proteinase and optionally a DNase inhibitor.

The methods of the invention are especially and most preferably useful for detecting DNA characteristics indicative of a subpopulation of transformed cells in a representative stool sample. The DNA characteristics may be, for example, mutations, including point mutations, deletions, additions, translocations, substitutions, and loss of  
25 heterozygosity. Methods of the invention may further comprise a visual examination of the colon. Finally, surgical resection of abnormal tissue may be done in order to prevent the spread of cancerous or precancerous tissue.

Accordingly, methods of the invention provide means for screening for the presence of a cancerous or precancerous subpopulation of cells in a heterogeneous  
30 sample, such as a stool sample. Methods of the invention reduce morbidity and mortality associated with lesions of the colonic epithelium. Moreover, methods of the

- 5 -

invention comprise more accurate and convenient screening methods than are currently available in the art, because such methods take advantage of the increased yield of relevant DNA.

Methods of the invention thus provide unexpected and enhanced detection and  
5 analysis of low-frequency DNA in a heterogeneous sample is facilitated through application of the methods described herein. That is, homogenization of stool sample in solvent at a ratio of at least 5:1 (volume to mass) alone, or in combination with methods for sample enrichment disclosed herein, provides a reliable method for obtaining a sufficient number of DNA molecules for effective and efficient analysis, even  
10 if the target molecule is a low-frequency DNA molecule. Further aspects and advantages of the invention are contained in the following detailed description thereof.

#### DESCRIPTION OF THE DRAWINGS

Figure 1 is a representation of a partial nucleotide sequence of the *kras* gene (base pairs 6282-6571) and the positions of capture probe CP1, PCR primer A1, and  
15 PCR primer B1, in relation to the *kras* nucleotide sequence.

Figure 2 is an image produced using a Stratagene Eagle Eye II Still Video System (Stratagene, La Jolla, CA), of the results of a gel electrophoresis run with the uncut DNA extracted as described in Example 2.

Figure 3 is an image produced using a Stratagene Eagle Eye II Still Video  
20 System (Stratagene, La Jolla, CA), of the results of a gel electrophoresis run with the DNA extracted as described in Example 3.

#### DETAILED DESCRIPTION OF THE INVENTION

The invention provides improved methods for extraction and analysis of nucleic acids from stool. According to methods of the invention, the yield of nucleic acids  
25 extracted from stool is increased by homogenizing the stool in a buffer at optimal ratio of buffer volume to stool mass. Yield is further improved by enriching for human DNA. Improved nucleic acid yields allow nucleic acid analysis of stool samples to be conducted more efficiently with less stool volume.

In preferred methods of the invention a stool sample obtained for analysis  
30 comprises at least a cross-section of a whole stool. As provided in U.S. Patent

- 6 -

No. 5,741,650, incorporated by reference herein, cells and cellular debris from the colonic epithelium is deposited onto and into stool in a longitudinal streak. Obtaining at least a cross-section of a stool ensures that a representative sampling of colonic epithelial cells and cellular debris is analyzed.

5        Once the stool sample is collected, it is homogenized in a physiologically acceptable solvent. A preferred means of homogenization employs agitation with glass beads. Physiologically acceptable solvents include those solvents generally known to those skilled in the art as suitable for dispersion of biological sample material. Such solvents include phosphate-buffered saline comprising a salt, such as 20-100mM NaCl  
10   or KCl, and optionally a detergent, such as 1-10% SDS or Triton™, and/or a proteinase, such as proteinase K (at, e.g., about 20mg/ml). A preferred solvent is a physiologically-compatible buffer comprising, for example, 1M Tris, 0.5M EDTA, 5M NaCl and water to a final concentration of 500mM Tris, 16mM EDTA and 10mM NaCl at pH 9. The buffer acts as a solvent to disperse the solid stool sample during  
15   homogenization. Applicants have discovered that increasing the volume of solvent in relation to solid mass of the sample results in increased yields of DNA.

      According to methods of the invention, solvent (buffer) is added to the solid sample in a solvent volume to solid mass ratio of at least about 5:1. The solvent volume to solid mass ratio is preferably in the range of about 10:1 to about 30:1, and  
20   more preferably in the range of about 10:1 to about 20:1. Most preferably, the solvent volume to solid mass ratio is about 10:1. Typically, solvent volume may be measured in milliliters, and solid mass measured in milligrams, but the practitioner will appreciate that the ratio of volume to mass remains constant, regardless of scale up or down of the particular mass and volume units. That is, solvent volume to solid mass ratios may be  
25   measured as liters:grams or  $\mu\text{l}$ :  $\mu\text{g}$ .

      In a preferred embodiment of the present invention, the homogenized sample is enriched for the target (human) DNA. In the context of the present invention, "enrichment" of the sample means manipulating the sample to decrease the amount of undesired, non-human DNA in the sample relative to the amount of target human DNA.  
30   Enrichment techniques include sequence-specific capture of target DNA or removal of bacterial nucleic acids.



- 7 -

In a preferred embodiment of the invention, the enrichment step is carried out in a physiologically compatible buffer, such as guanidine isothiocyanate (GITC). Capture probes are then added to the mixture to hybridize to target DNA in order to facilitate selective removal of target DNA from the sample.

5        Sequence specific capture of target DNA can be accomplished by initially denaturing sample DNA to form single-stranded DNA. Then, a sufficient quantity of sequence specific oligonucleotide probe that is complementary to at least a portion of a target polynucleotide (e.g., a sequence in or near the p53 allele) is added. The probe sequence (labeled with biotin) is allowed to hybridize to the complementary target DNA  
10        sequence. Beads coated with avidin or streptavidin are then added and attach to the biotinylated hybrids by affinity-binding. The beads may be magnetized to facilitate isolation.

After separation of probe-target hybrids, the resultant DNA is washed repeatedly to remove inhibitors, including those commonly introduced via the capture probe  
15        technique. In the methods of the present invention, washes are preferably carried out approximately four times with 1M GITC and 0.1% detergent, such as Igepal (Sigma). The initial washes are then preferably followed by two washes with a standard wash buffer (such as Tris-EDTA-NaCl) to remove the GITC from the mix, since GITC is a known inhibitor of DNA polymerases, including those associated with PCR.

20        Finally, the target DNA is eluted into a small volume of distilled water by heating. Assays using polymerase chain reaction (PCR), restriction fragment length polymorphism (RFLP) analysis or other nucleic acid analysis methods may be used to detect DNA characteristics indicative of a disorder, such as colorectal cancer or pre-cancer. Several particularly useful analytical techniques are described in co-pending  
25        applications Serial Number 08/700,583, 08/815,576 and 08/877,333, the disclosures of which are incorporated herein by reference.

In an alternative embodiment, the homogenized sample is examined to determine that the sample has at least a minimum number (N) of total DNA molecules to provide for detection of a low-frequency target DNA molecule. The number of  
30        molecules analyzed in a sample determines the ability of the analysis to detect low-frequency events. In the case of PCR, the number of input molecules must be about

- 8 -

500 if the PCR efficiency is close to 100%. As PCR efficiency goes down, the required number of input molecules goes up. Analyzing the minimum number of input molecules reduces the probability that a low-frequency event is not detected in PCR because it is not amplified in the first few rounds. Methods of the invention therefore include

5 determining a threshold number of sample molecules that must be analyzed in order to detect a low-frequency molecular event at a prescribed level of confidence.

As is more fully described in copending Application Serial No. \_\_\_\_\_ [Atty Docket No. EXT-021], which is incorporated herein by reference, the determination of a minimum number N of DNA molecules that must be present in a

10 sample to permit amplification and analysis of a low-frequency target DNA molecule is based upon a model of stochastic processes in PCR. Utilizing pre-set or predetermined values for PCR efficiency and mutant DNA to wild-type DNA ratio in the sample, the model predicts the number of molecules that must be presented to the PCR in order to ensure, within a defined level of statistical confidence, that a low-frequency molecule

15 will be amplified.

The skilled practitioner will appreciate that determination of the minimum number N of molecules present in the sample may be used in lieu of, or in addition to, the enrichment techniques detailed above, to ensure reliable results in the methods of the present invention.

20 Alternatively, methods of the invention may also be used to isolate total DNA from stool homogenate. The homogenized mixture is centrifuged to form a pellet made up of cell debris and stool matter, and a supernatant containing nucleic acid and associated proteins, lipids, etc. The supernatant is treated with a detergent, such as 20% SDS, and enzymes capable of degrading protein (e.g., Proteinase K). The

25 supernatant is then Phenol-Chloroform extracted. The resulting purified nucleic acids are then precipitated by means known in the art. A variety of techniques in the art can then be employed to manipulate the resulting nucleic acids, including further purification or isolation of specific nucleic acids.

Methods of the invention are also useful for analysis of pooled DNA samples. As

30 described in more detail in Application Serial No. 09/098,180, and U.S. Patent No. 5,670,325, both of which are incorporated by reference herein, enumerative

- 9 -

analysis of pooled genomic DNA samples is used to determine the presence or likelihood of disease. Pooled genomic DNA from healthy members of a population and pooled genomic DNA from diseased members of a population are obtained. The number or amount of each variant at a single-nucleotide polymorphic site is determined  
5 in each sample. The numbers or amounts are analyzed to determine if there is a statistically-significant difference between the variant(s) present in the sample obtained from the healthy population and those present in the sample obtained from the diseased population. A statistically-significant difference indicates that the polymorphic locus is a marker for disease.

10 These methods may be used to identify a nucleic acid (e.g., a polymorphic variant) associated with a disease. Such methods comprise counting the number or determining the amount of a nucleic acid, preferably a single base, in members of a diseased population, and counting numbers or determining amounts of the same nucleic acid in members of a healthy population. A statistically-significant difference in  
15 the numbers of the nucleic acid between the two populations is indicative that the interrogated locus is associated with a disease.

Once the polymorphic locus is identified, either by methods of the invention or by consulting an appropriate database, such methods are useful to determine which variant at the polymorphic locus is associated with a disease. In this case, enumerative  
20 methods are used to determine whether there is a statistically-significant difference between the number of a first variant in members of a diseased population, and the number of a second variant at the same locus in members of a healthy population. A statistically-significant difference is indicative that the variant in members of the diseased population is useful as a marker for disease. Using this information, patients  
25 are screened for the presence of the variant that is thought to be associated with disease, the presence such a variant being indicative of the presence of disease, or a predisposition for a disease.

Methods of the present invention are particularly useful for isolation and analysis of nucleic acids that encompass genes that have mutations implicated in colorectal  
30 cancer, such as *kras*. The *kras* gene has a length of more than 30 kbp and codes for a 189 amino acid protein characterized as a low-molecular weight GTP-binding protein.

- 10 -

The gene acquires malignant properties by single point mutations, the most common of which occurs at the 12th amino acid. Several studies have confirmed that approximately 40% of primary colorectal adenocarcinoma cells in humans contain a mutated form of the kras gene. Accordingly, the kras gene is a particularly suitable target for the methods of colorectal cancer detection of the present invention.

Toward this end, applicants have constructed a suitable exemplary capture probe directed to the kras nucleotide sequence. The capture probe, designated CP1, has the following sequence: 5' GCC TGC TGA AAA TGA CTG AAT ATA AAC TTG TGG TAG T 3' (SEQ. ID NO: 1), and is preferably biotinylated at the 5' end in order to facilitate isolation. As illustrated more fully below, CP1 is effective in the sequence specific capture of kras DNA.

Suitable PCR primers for the analysis of extracted kras DNA sequence have also been determined. Primer A1 has the sequence: 5' C CTG CTG AAA ATG ACT GAA 3' (SEQ ID NO: 2), and Primer B1 has the sequence: 5' CAT GAA AAT GGT CAG AGA AA 3' (SEQ ID NO: 3). The PCR primers A1 and B1, as well as capture probe CP1, are depicted in Figure 1, showing their relation to the kras nucleotide sequence, base pairs 6282-6571 (SEQ ID NO: 4). One skilled in the art can construct other suitable capture probes and PCR primers for kras or other target genes or nucleotide sequences, using techniques well known in the art.

Accordingly, the methods of the present invention, which involve homogenizing stool sample in a volume of solvent such that the ratio of solvent volume to stool mass is at least 5:1, and/or enriching the sample for human DNA, provide a means for obtaining a sample having a minimum number N of total DNA molecules to facilitate detection of a low-frequency target DNA molecule. These methods thus provide the unexpected result that one is now able to reliably detect a small portion of low-frequency DNA in a heterogeneous sample.

The following examples provide further details of methods according to the invention. However, numerous additional aspects of the invention will become apparent upon consideration of the following examples.

- 11 -

**Example 1****Stool Sample Preparation**

Voided stool was collected from a patient and a cross-sectional portion of the stool was removed for use as a sample. After determining the mass of the sample, an  
5 approximately 10x volume of Tris-EDTA-NaCl lysis buffer was added to the solid sample in a test tube. The final concentration of the buffer was 500mM Tris, 16mM EDTA and 10mM NaCl, at a pH of about 9.0. Four 10mm glass balls were placed in the tube and the tube and contents were homogenized in an Exactor II shaker for 15 minutes. The homogenized mixture was then allowed to stand 5 minutes at room  
10 temperature. The tube was then centrifuged for 5 minutes at 10,000 rpm in a Sorvall Centrifuge, and the supernatant was transferred to a clean test tube. A 20% SDS solution was added to the tube to a final concentration of 0.5%. Proteinase K was also added to the tube to a final concentration of 500mg/ml. The tube was then incubated overnight at 37°C.

15 After incubation, the contents of the tube were extracted with an equal volume of phenol/chloroform and centrifuged at 3500 rpm for 3 minutes. The aqueous layer was then transferred to a new tube and extracted three (3) times with equal volumes of chloroform and centrifuged at 3500 rpm for 3 minutes. The aqueous layer was then transferred to a new tube and 0.1x volume of 3M NaOAc was added to the aqueous  
20 portion, which was then extracted with an equal volume of isopropanol, and centrifuged for 5 minutes at 12,000 rpm. The supernatant was discarded, and the pellet was washed with 10ml of 70% ethanol, and centrifuged at 12,000 rpm for 5 minutes. The supernatant was discarded and the pellet containing isolated DNA was dried by inverting the tube.

**25 Example 2**

A comparative analysis of solvent volume to mass ratios was conducted. Three separate stool samples were prepared as described above. A first sample, designated SS88-3x, was homogenized in buffer at a volume to mass ratio of 3:1. A second sample, designated SS88-5x, was homogenized at a ratio of 5:1; and a third sample,  
30 designated SS88-10x, was homogenized at a ratio of 10:1.

- 12 -

Total DNA from each sample was resuspended in 100  $\mu$ l of 100 mM Tris, 10 mM EDTA buffer and 10  $\mu$ l aliquots were loaded onto a 4% agarose gel for electrophoresis at 125 V constant voltage for about one hour. The results are shown in Figure 2. As shown in Figure 2, the yield of total DNA increased as the ratio of solvent to mass  
5 increased from 3x to 10x.

### Example 3

A second set of four equivalent samples was prepared from a single stool sample. Each of the four samples was of equal mass, and was homogenized as described in Example 1 at a solvent volume to stool mass ratio of 5:1, 10:1, 20:1, and  
10 30:1, respectively. After homogenization each sample was subdivided into 8 aliquots, 4 treated with RNase, and 4 untreated. Total DNA was then isolated as described above and analyzed on agarose gels.

The results are shown in Figure 3. As shown, a ratio of 10:1 produced the greatest yield of nucleic acids. Figure 3 also shows the effect of RNase treatment on  
15 the yield of DNA from each stool sample. As shown in the Figure, RNase treatment virtually eliminates RNA from the sample, but leaves DNA intact. The results indicate that optimal solvent volume to stool mass ratios greatly increase DNA yield from stool samples.

### Example 4

20 Sequence-Specific Capture of target DNA.

Once extracted from stool, specific nucleic acids are isolated using sequence-specific capture probes. Total DNA was extracted from a stool sample according to the methods described in Example 1. The pelletized DNA was resuspended in 1ml of TE buffer. A 100  $\mu$ l aliquot of this solution was removed to a new tube and 100  $\mu$ l of 6M  
25 guanidine isothiocyanate (GITC) was added to a final concentration of 3M GITC. A vast excess of biotinylated kras capture probe CP1 was added to the sample. The mixture was heated to 95°C for 5 minutes to denature the DNA, then cooled to 37°C for 5 minutes. Finally, probe and target DNA were allowed to hybridize for 30 minutes at room temperature. Streptavidin-coated magnetized beads (320 mg) (Dynal Corp.) were  
30 suspended in 400  $\mu$ l distilled water and added to the mixture. After briefly mixing, the tube was maintained at room temperature for 30 minutes.

- 13 -

Once the affinity binding was completed, a magnetic field was applied to the sample to draw the magnetized isolation beads (both with and without hybridized complex out of the sample. The beads were then washed four (4) times in 1M GITC/0.1% Igepal (Sigma, St. Louis, MO) solution for 15 minutes, followed by two (2) washes with wash buffer (TE with 1M NaCl) for 15 minutes in order to isolate complexed streptavidin. Finally, 10 µl distilled water was added to the beads and heated at 95°C for 3 minutes to elute the DNA. Sequencing and/or gel electrophoresis enable confirmation of the capture of kras-specific DNA.

Accordingly, methods of the invention produce increased yields of DNA from stool, thereby allowing more efficient sequence-specific capture of target nucleic acid. Methods of the invention provide improvements in the ability to detect disease-related nucleic acid mutations present in stool. The skilled artisan will find additional applications and embodiments of the invention useful upon inspection of the foregoing description of the invention. Therefore, the invention is limited only by the scope of the appended claims.

- 14 -

**CLAIMS**

What is claimed is:

- 1 1. A method for analyzing DNA extracted from stool, comprising:
  - 2 homogenizing a stool sample in a solvent for DNA in order to form a
  - 3 homogenized sample mixture having a solvent volume to stool mass ratio of at
  - 4 least 5:1;
  - 5 enriching said homogenized sample for human DNA; and
  - 6 analyzing said human DNA for characteristics of disease.
- 1 2. The method of claim 1 wherein the solvent volume to stool mass ratio is from  
2 about 10:1 to about 30:1.
- 1 3. The method of claim 2 wherein the solvent volume to stool mass ratio is about  
2 10:1 to about 20:1.
- 1 4. The method of claim 2 wherein the solvent volume to stool mass ratio is about  
2 10:1.
- 1 5. The method of claim 1 wherein the solvent comprises a physiologically  
2 compatible buffer.
- 1 6. The method of claim 5 wherein the buffer comprises Tris-EDTA-NaCl.
- 1 7. The method of claim 6 wherein the Tris-EDTA-NaCl buffer comprises a final  
2 concentration of about 50mM Tris, about 16 mM EDTA and about 10mM NaCl at  
3 about pH 9.0.
- 1 8. The method of claim 1 wherein the solvent comprises guanidine isothiocyanate  
2 buffer.
- 1 9. The method of claim 8 wherein the guanidine isothiocyanate buffer comprises a  
2 final concentration of from about 1 to about 5 M.



- 15 -

- 1 10. The method of claim 9 wherein the guanidine isothiocyanate buffer comprises a  
2 final concentration of about 3 M.
- 1 11. The method of claim 1 wherein said enriching step comprises contacting said  
2 DNA with a sequence-specific capture probe.
- 1 12. The method of claim 1 wherein said solvent comprises a detergent and a  
2 proteinase.
- 1 13. The method of claim 1 wherein said DNA is human DNA.
- 1 14. A method of screening for the presence of a colorectal cancerous or pre-  
2 cancerous lesion in a patient, the method comprising the steps of:
- 3 obtaining a sample comprising at least a cross-sectional portion of a stool voided by the  
4 patient;
- 5 homogenizing the sample in a solvent in order to form a homogenized sample mixture  
6 having a solvent volume to stool mass ratio of at least 5:1;
- 7 enriching said sample for a target human DNA; and
- 8 analyzing the target human DNA for DNA characteristics indicative of the presence of  
9 said colorectal cancerous or pre-cancerous lesion.
- 1 15. The method of claim 14 wherein said analyzing step comprises amplifying the  
2 DNA with a polymerase chain reaction.
- 1 16. The method of claim 14 wherein said DNA characteristics comprise a loss of  
2 heterozygosity encompassing a polymorphic locus.
- 1 17. The method of claim 14 wherein said DNA characteristic is a mutation.
- 1 18. The method of claim 17 wherein said mutation is selected from the group  
2 consisting of loss of heterozygosity and microsatellite instability.

- 16 -

1 19. The method of claim 14 wherein said DNA characteristics comprise a deletion in  
2 a tumor suppressor allele.

1 20. The method of claim 14 wherein said analyzing step comprises determining  
2 whether a difference exists in said sample between a number X of a first allele  
3 known or suspected to be mutated in a subpopulation of cells in the sample and  
4 a number Y of a second allele that is known or suspected not to be mutated in a  
5 subpopulation of cells in the sample, the presence of a statistically-significant  
6 difference being indicative of a mutation in a subpopulation of cells in the sample  
7 and the potential presence of a cancerous or precancerous lesion.

1 21. The method of claim 14 wherein said analyzing step comprises determining  
2 whether a difference exists between a number of a target tumor suppressor  
3 allele in the sample and a number of a non-cancer-associated reference allele in  
4 the sample, the presence of a statistically-significant difference being indicative  
5 of a deletion of the target tumor suppressor allele in a subpopulation of cells in  
6 the sample and the potential presence of a cancerous or precancerous lesion.

1 22. The method of claim 14 wherein said analyzing step further comprises the steps  
2 of:

3 a) detecting an amount of a maternal allele at a polymorphic locus in the  
4 sample;

5 b) detecting an amount of a paternal allele at the polymorphic locus in the  
6 sample; and

7 c) determining whether a difference exists between the amounts of  
8 maternal and paternal allele,

9 the presence of a statistically-significant difference being indicative of a deletion  
10 at the polymorphic locus in a subpopulation of cells in the sample and the potential  
11 presence of a lesion.

- 17 -

1 23. The method of claim 22 wherein said polymorphic locus is a single base  
2 polymorphism and is heterozygous between said maternal and paternal alleles.

1 24. The method of claim 22 wherein said detecting steps comprise,

- 2 a) hybridizing probe to a portion of said polymorphic locus on both  
3 maternal and paternal alleles that is immediately adjacent to said  
4 single-base polymorphism;
- 5 b) exposing said sample to a mixture of detectably-labeled dideoxy  
6 nucleoside triphosphates under conditions which allow appropriate  
7 binding of said dideoxy nucleoside triphosphates to said single-  
8 base polymorphism;
- 9 c) washing the sample; and
- 10 d) counting an amount of each detectably-labeled dideoxy nucleoside  
11 triphosphate remaining for the sample.

1 25. The method of claim 24 wherein said detectable label is selected from the group  
2 consisting of radioisotopes, fluorescent compounds, and particles.

1 26. The method of claim 14 wherein said analyzing step comprises a method for  
2 detecting heterozygosity at a single-nucleotide polymorphic locus, comprising the steps  
3 of:

- 4 a) hybridizing probes to a sequence immediately adjacent to a single-  
5 base polymorphism;
- 6 b) exposing the sample to a plurality of different labeled dideoxy  
7 nucleotides
- 8 c) washing the sample;
- 9 d) determining which of said dideoxy nucleotides are incorporated into  
10 said probes; and

- 18 -

11 e) detecting heterozygosity at the single-nucleotide polymorphic site  
12 as the detection of two dideoxy nucleotides having been  
13 incorporated into the probe.

1 27. The method of claim 14 wherein said analyzing step comprises:

2 (a) exposing the sample to a plurality of a first oligonucleotide probe  
3 and to a plurality of a second oligonucleotide probe under hybridization conditions,  
4 thereby to hybridize

5 (1) said first oligonucleotide probes to copies of a first  
6 polynucleotide segment characteristic of wild-type cells of the organism, and

7 (2) said second oligonucleotide probes to copies of a second  
8 polynucleotide segment characteristic of a wild-type genomic region suspected to be  
9 deleted or mutated in colorectal cancer cells;

10 (b) detecting a first number of duplexes formed between said first  
11 probe and said first segment and a second number of duplexes formed between said  
12 second probe and said second segment; and

13 (c) determining whether there is a difference between the number of  
14 duplexes formed between said first probe and said first segment and the number of  
15 duplexes formed between said second probe and said second segment,

16 the presence of a statistically-significant difference being indicative of the  
17 presence in said sample of a colorectal cancer or precancerous lesion.

1 28. The method of claim 27 wherein said first and second oligonucleotide probes  
2 each are coupled to a distinct detectable label.

1 29. The method of claim 27 wherein

2 said first oligonucleotide probes are attached to a first particle in a ratio of  
3 one first oligonucleotide probe to one particle and said second oligonucleotide probes

- 19 -

4 are attached to a second particle detectably distinct from said first particle in a ratio of  
5 one second oligonucleotide probe to one second particle, wherein

6 said detecting step comprises separating hybridized from unhybridized  
7 first and second oligonucleotide probes and subsequently passing hybridized first and  
8 second oligonucleotide probes through a detector to determine said first and second  
9 numbers.

1 30. The method of claim 29 wherein said first and second particles are of detectably  
2 different sizes.

1 31. The method of claim 29 wherein said first and second particles are of detectably  
2 different colors.

1 32. The method of claim 27 further comprising, prior to step a) the steps of  
2 converting double-stranded DNA in said sample to single-stranded DNA and removing  
3 complement to said first and second polynucleotide segments.

1 33. The method of claim 32 wherein said removing step comprises hybridizing said  
2 complement to a nucleic acid probe attached to a magnetic particle and subsequently  
3 removing said magnetic particle from the sample.

1 34. The method of claim 14 wherein said analyzing step comprises a method for  
2 detecting a nucleic acid sequence change in a target allele in the sample, comprising  
3 the steps of:

4 (a) determining

5 (i) an amount of wild-type target allele in the sample, and

6 (ii) an amount of a reference allele in the sample; and

7 (b) detecting a nucleic acid sequence change in the target allele in the  
8 sample,

- 20 -

9                   a statistically significant difference in the amount wild-type target allele  
10                and the amount of reference allele obtained in said determining step being  
11                indicative of a nucleic acid sequence change.

1    35.   The method according to claim 34 wherein said determining step comprises  
2    exposing said sample to a first oligonucleotide probe capable of hybridizing with a  
3    portion of said wild-type allele and to a second oligonucleotide probe capable of  
4    hybridizing to a portion of said reference allele, and removing from said sample any  
5    unhybridized first or second oligonucleotide probe.

1    36.   A method for screening for the presence of a colorectal cancerous or  
2    precancerous lesion in a patient, the method comprising the steps of:

3                obtaining a sample comprising at least a cross-sectional portion of a stool voided  
4    by the patient;

5                homogenizing the sample in a solvent in order to form a homogenized sample  
6    mixture having a solvent volume to stool mass ratio of at least 5:1;

7                ensuring that said sample have at least a minimum number N of total DNA  
8    molecules to provide for detection of a low-frequency target DNA molecule;

9                analyzing the target DNA for DNA characteristics indicative of the presence of  
10   said colorectal cancerous or pre-cancerous lesion.

1    37.   A method for screening for the presence of a colorectal cancerous or  
2    precancerous lesion in a patient, the method comprising the steps of:

3                obtaining a sample containing of least a cross-sectional portion of a stool voided  
4    by the patient;

5                homogenizing the sample in a solvent in order to form a homogenized sample  
6    mixture having a solvent volume to stool mass ratio of at least 5:1;

7                enriching said homogenized sample for target human DNA;

- 21 -

8           ensuring that said sample have at least a minimum number N of total DNA  
9           molecules to provide for detection of a low-frequency target DNA molecule;

10          analyzing the target human DNA for DNA characteristics indicative of the  
11          presence of said colorectal cancerous or pre-cancerous lesion.

1    38.    The method of claim 36 wherein said analyzing step comprises amplifying the  
2    DNA with a polymerase chain reaction.

1    39.    The method of claim 37 wherein said analyzing step comprises amplifying the  
2    DNA with a polymerase chain reaction.

1    40.    The method of claim 37 wherein said enriching step comprises contacting said  
2    DNA with a sequence-specific capture probe.

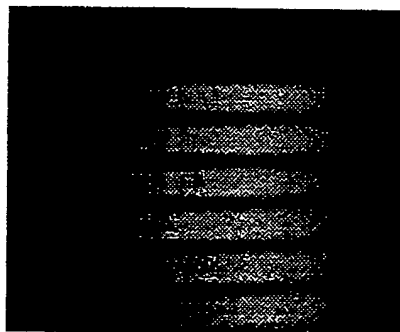
5' -GT ACT GGT GGA GTA TTT GAT AGT GTA TTA ACC TTA  
 TGT GTG ACA TGT TCT AAT ATA GTC ACA TTT TCA TTA TTT  
 TTA TTA TAA GGC CTG CTG AAA ATG ACT GAA TAT AAA  
 CTT GTG GTA GTT GGA GCT GGT GGC GTA GGC AAG AGT GCC  
 TTG ACG ATA CAG CTA ATT CAG AAT CAT TTT GTG GAC GAA  
 TAT GAT CCA ACA ATA GAG GTA AAT CTT GTT TTA ATA  
 TGC ATA TTA CTG GTG CAG GAC CAT TCT TTG ATA CAG ATA  
 AAG GTT TCT CTG ACC ATT TTC ATG TAC AGA AGT CCT TGC  
 TAA GA-3'

CP1  
 A1  
 B1

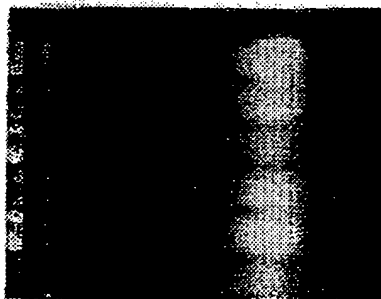
FIG. 1



SS88-10x  
m n o p q r



SS88-5x  
g h i j k l



SS88-3x  
a b c d e f

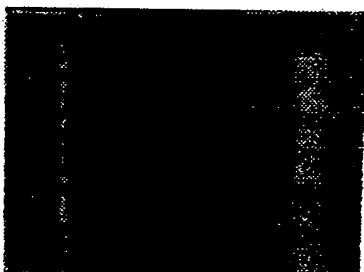
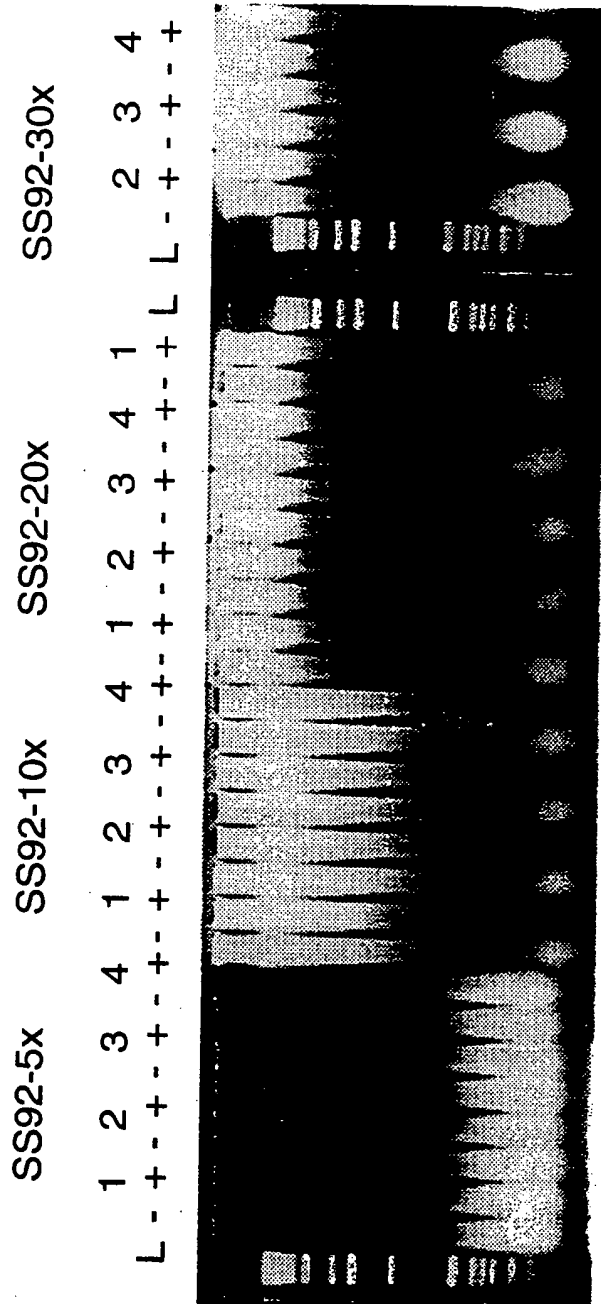


FIG. 2



**FIG. 3**

## SEQUENCE LISTING

<110> Shuber, Anthony P  
Lapidus, Stanley M  
Radcliffe, Gail E

<120> Methods for Stool Sample Preparation

<130> EXT-028PC

<140>

<141>

<150> US 09/198,083

<151> 1998-11-23

<160> 4

<170> PatentIn Ver. 2.0

<210> 1

<211> 37

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Capture probe  
C1

<400> 1

gcctgctgaa aatgactgaa tataaacttg tggtagt

37

<210> 2

<211> 19

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: PCR primer A1

<400> 2

cctgctgaaa atgactgaa

19

<210> 3

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: PCR primer B1

<400> 3

catgaaaatg gtcagagaaa

20

<210> 4

<211> 307

<212> DNA

<213> Homo sapiens

<220>

<223> Partial nucleotide sequence of the kras gene

<400> 4

gtactgggtgg agtatttgat agtgatttaa ccttatgtgt gacatgttct aatatagtca 60  
cattttcatt atttttatta taaggcctgc tgaaaatgac tgaatataaa cttgtggtag 120  
ttggagctgg tggcgtaggc aagagtgccct tgacgataca gctaattcag aatcattttg 180  
tggacgaata tgatccaaca atagaggtaa atcctgtttt aatatgcata ttactgggtgc 240  
aggaccattc tttgatacag ataaaggttt ctctgaccat tttcatgtac agaagtcctt 300  
gctaaga

307